CITY OF ROCHESTER

Team 5

Brynn Lee, Eugene Ayonga, Homayra Tabassum,
Medhini Sridharr, Nour Assili

Goals

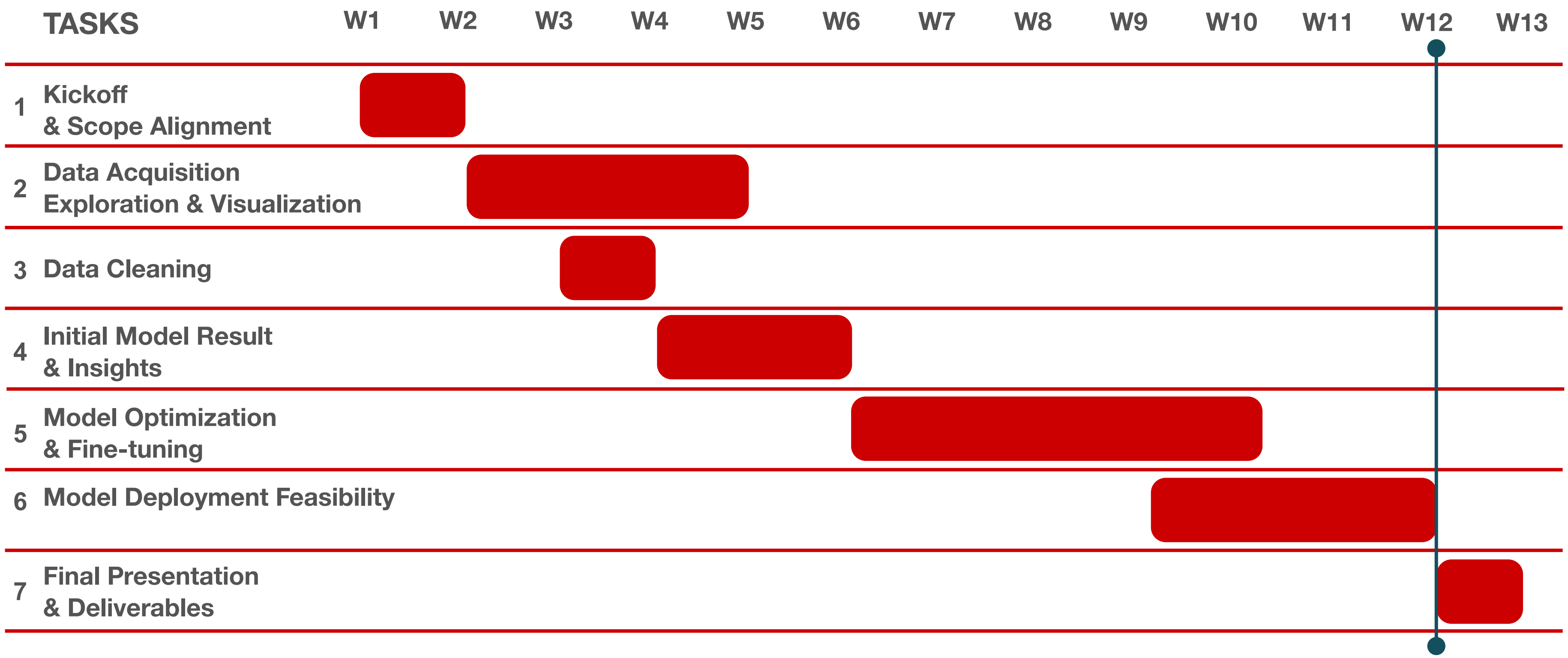
1. Reduce response times

Analyze historical incident and unit data to identify patterns and optimize personnel and equipment allocation.

2. Assess the need for new fire stations

Define and evaluate factors, ensuring the fire department can meet future demands. If a new station is needed, recommend optimal locations.

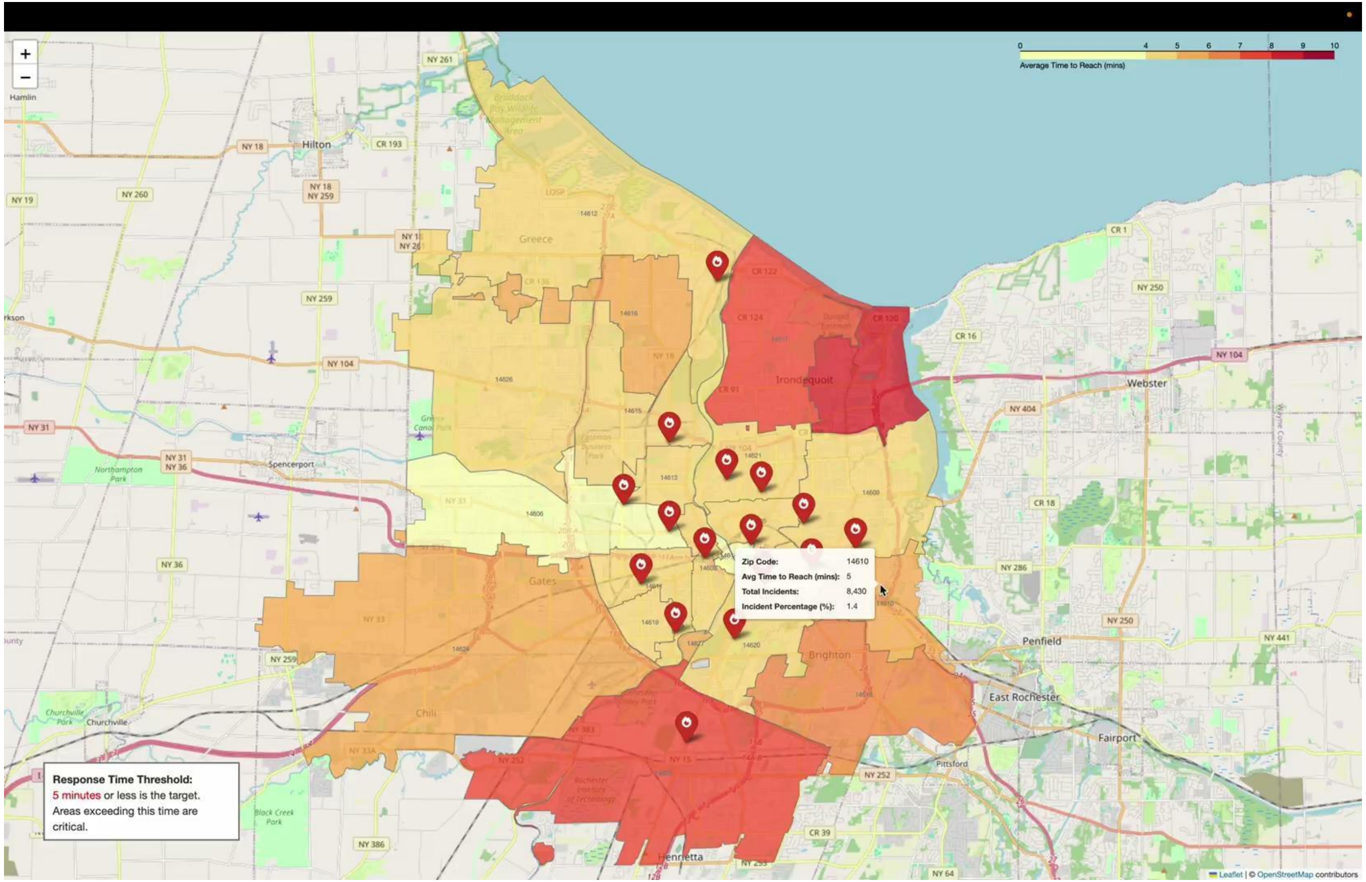
Milestones


TASKS
W1
W2
W3
W4
W5
W6
W7
W8
W9
W10
W11
W12
W13
1 Kickoff & Scope Alignment
2 Data Acquisition Exploration & Visualization
3 Data Cleaning
4 Initial Model Result & Insights
5 Model Optimization & Fine-tuning
6 Model Deployment Feasibility
7 Final Presentation & Deliverables

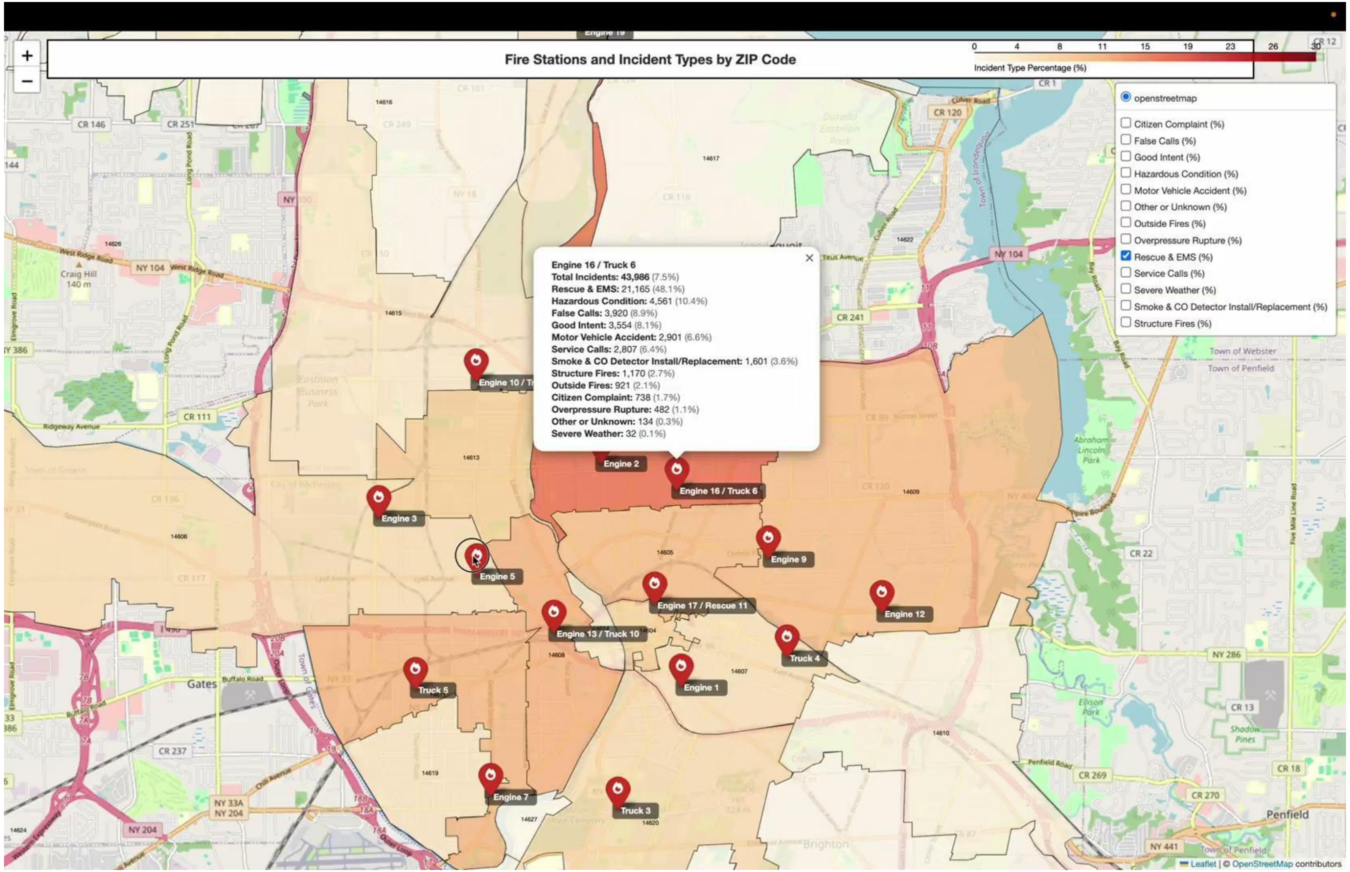
Data

Type	Description	Shape	Period
Unit Response 🚒	Tracks **units** (vehicles & personnel) - Logs unit-level responses - Key timestamps - Geographic, operational, and date/time data	(1M, 40) 57K/year	2008 - 2019
Incident 🚨	Tracks **incidents** - Logs incident type, situation found - Details: duration, personnel, action taken - Includes geographic and date/time data	(650K, 180) 37K/year	2006 - 2024
Station Apparatus 🧰	Facility ID, units, specialty units, descriptions	(16, 6)	-
Geospatial 🗺️	Geographic shapefile data for RFD locations	(16, 13)	-

Interactive Map: Time to Reach



Interactive Map: Incident Types



Resource Allocation - Suggested Changes

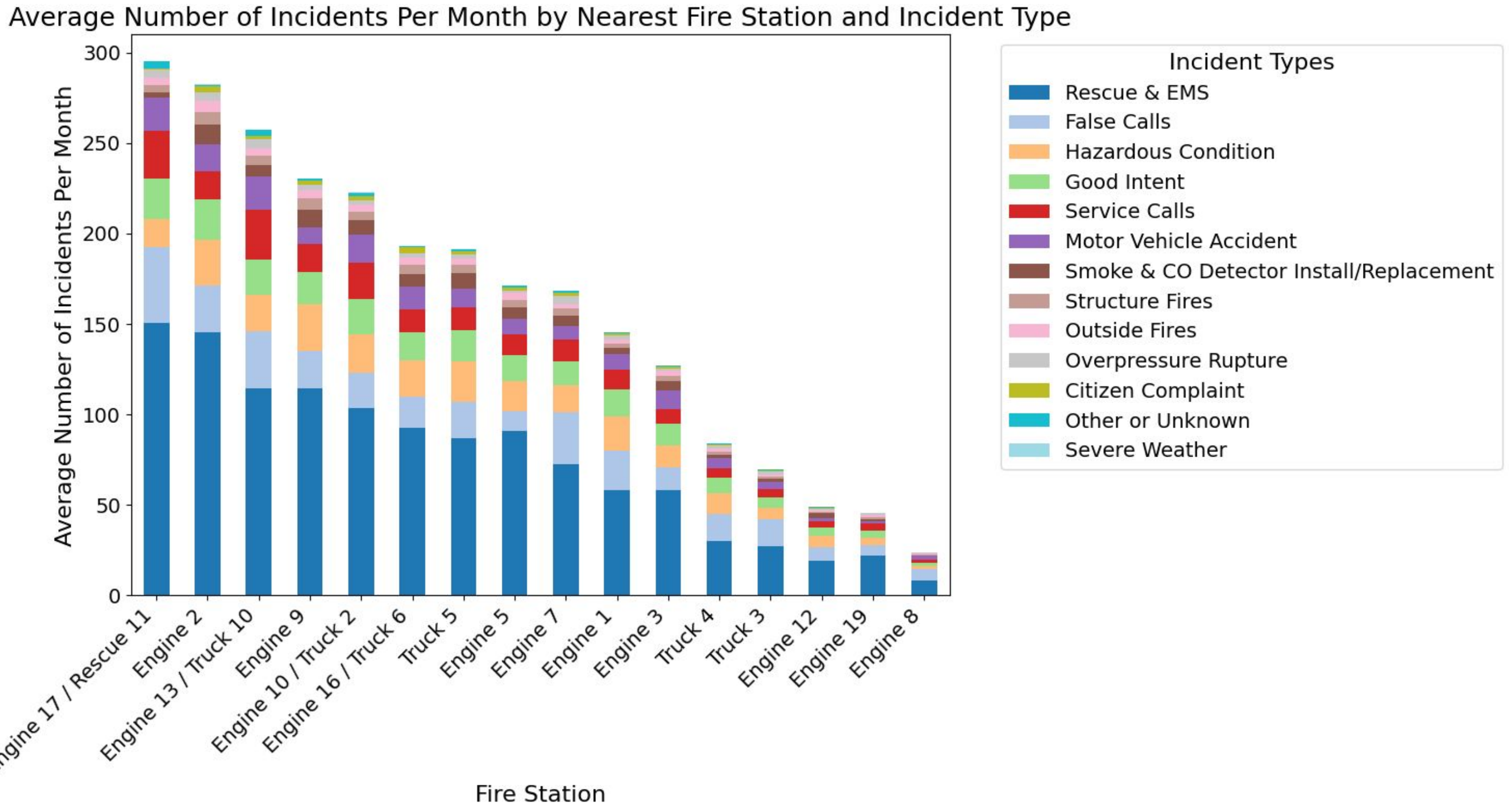
- **Unit E17 (Hazardous Condition)**:
  - Currently in **Engine 17/Rescue 11**
  - **Geo-spatial analysis** shows most "Hazardous Condition" incidents near **Engine 2**
  - Suggested exchange:
    - Move **E17** to **Engine 2**
    - Move **E2** to **Engine 17/Rescue 11**

- **Fire Investigation Units (CAR 91 – CAR 94, CAR 98)**:
  - Currently all 5 vehicles are at **Engine 10/Truck 2**
  - 80% of incidents handled are **"Service Calls"**
  - High "Service Calls" near **Engine 13/Truck 10**
  - Suggested change: Move **2 vehicles** to **Engine 13/Truck 10**

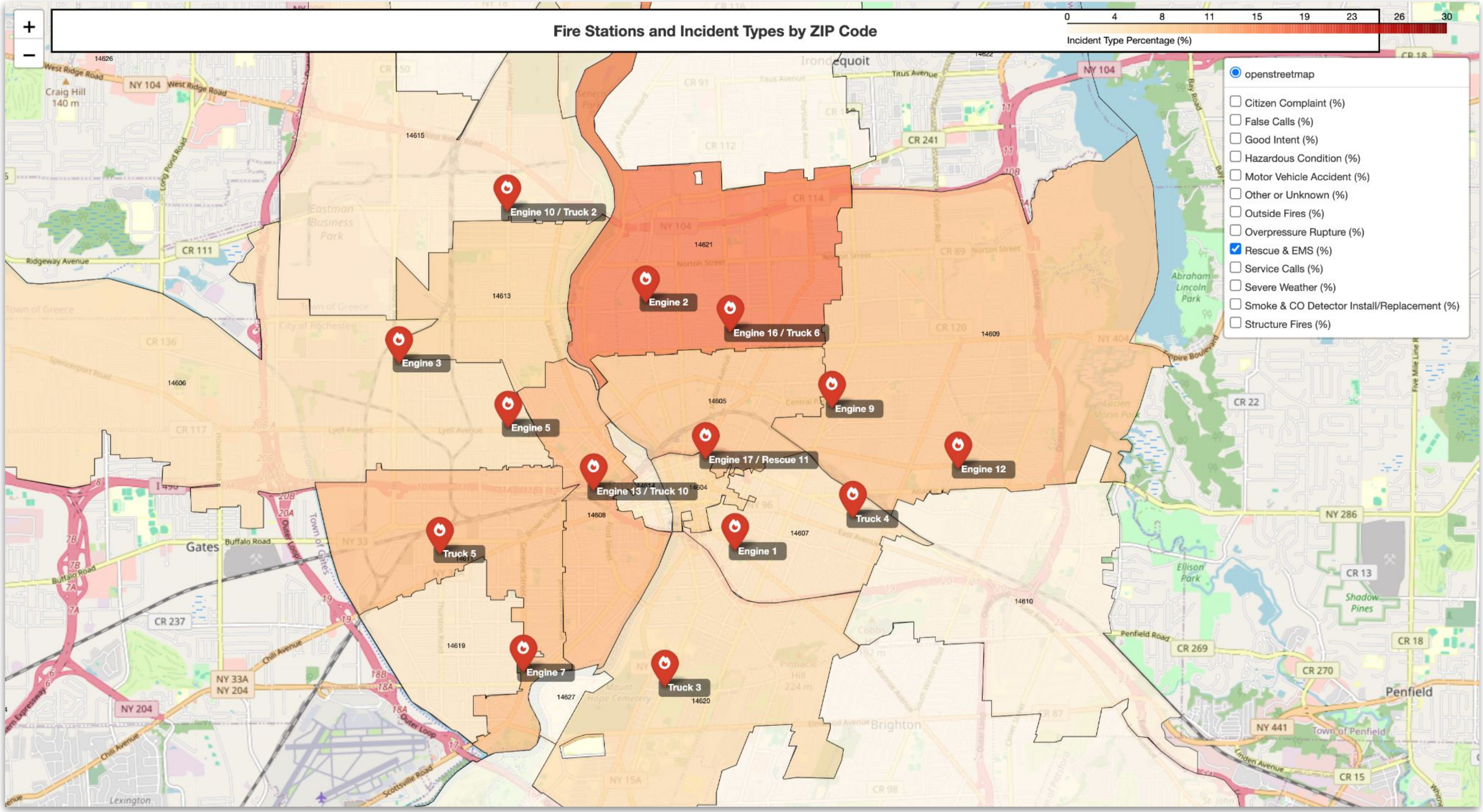
Overall Allocation Insights

- **Comments on Special Trucks/Engines**
  - **R11 (Rescue Truck)**:
    - Only rescue truck in the city
    - Located at **Engine 17/Rescue 11** (center of the city)
    - Current location is **optimal**
  - **E1 & E10 (Jaws of Life)**:
    - Special tools for vehicle rescues
    - E1: Located at **Engine 1** (one side of the city)
    - E10: Located at **Engine 10/Truck 2** (opposite side)
    - Current locations **cover different sides** of the city effectively
- **Overall Allocation**
  - Most units handle incidents **near their assigned Fire Station**
  - No significant patterns requiring relocation for other units
  - **Current allocation** is generally **effective**

Proportion of Incident Types Per Fire Station



High EMS Geographic Distribution



Proposed Low-Acuity Program

- **“Ridiculous” calls**: low-acuity and high frequency
- **Proposed program**: 2 firefighters and 1 social worker for non emergency EMS calls operating during peak call times for only 4 days per week
- **Objective**: To free up firefighters and police resources for what aligns with their training
- **Estimated funding**: $400,000

Example : Health One, a Seattle Fire Department program that responds to low-acuity crisis calls

Predicting Monthly Incident Counts

Step 1

Data Preparation

- Monthly Aggregation, Lag Feature Creation, Feature Engineering

Step 2

Feature Selection

- Random Forest + SHAP, Avg. R^2 = 85.28% across folds

Step 3

10-Year Feature Forecasting

- SARIMAX, Poisson Distrib., ETS Model, Covid-19 Impact, MSE, MAE, MAPE, AIC, BIC

Step 4

10-Year Incident Count Forecasting

- Prophet, RMSE, MAE, MAPE metrics

1. Monthly Incidents - Data Creation

Incident-Level Data Overview | Aggregation Process | Monthly Data Creation

Individual incidents from 2006 to 2024 for 16 RFD stations.

Key Variables:

- **Categorical:** Severity of incidents.
- **Geographical:** Incidents' latitude and longitude
- Other related incident details.

Numerical Variables: Aggregated using **mean** or **median** (based on distribution).

Categorical Variables:

- Applied **one-hot encoding.**
- Counted occurrences for each category by month-year per station.

Aggregated data to create **monthly incident counts** per fire station:

- Final output: 16 Month-year-wise structured datasets.
- Added lags in each dataset by looking at ACF plots of monthly incident counts

2. Feature Selection for Each Station

Multicollinearity

Features with correlation > 0.7.

Dropped features: Domain Knowledge and Correlation with Target

Created Interaction Terms:

- latitude * longitude (location interaction)
- civilian_deaths * alarm level

Feature Importance

Random Forest (RF) Analysis:

Ran Random Forest model with TSCV to identify top features contributing to predictive performance.

SHAP Analysis:

Based on RF model and extracted **top 10 impactful features** for each fire station.

Final Feature Selection

Overlapping Features:

Compared top features from Random Forest feature importances and SHAP top 10 impactful features.

Retained overlapping features as the **final set of important features** for each station.

Avg. R2 across models for each station: 85.28%

3. Effect of COVID-19

Split Data into Two Groups

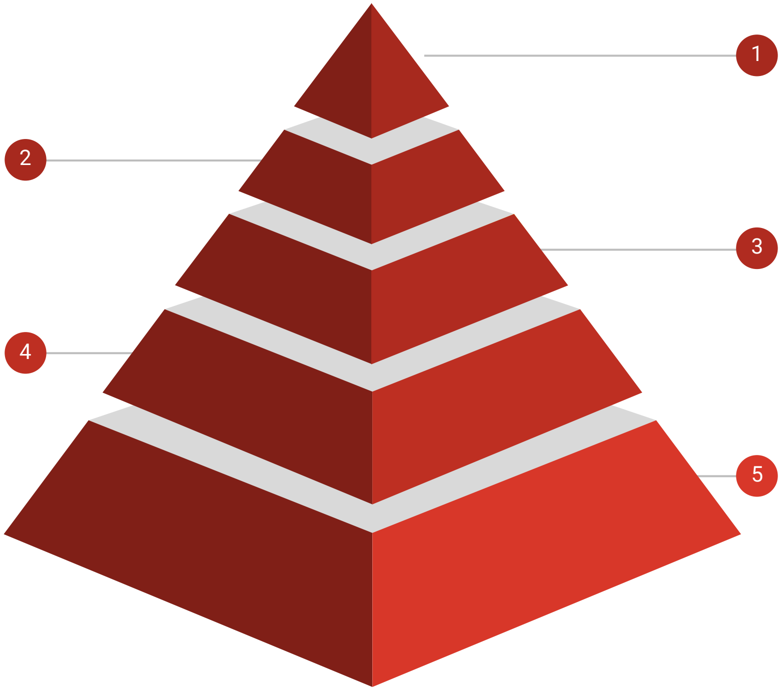
For each feature: Split into Group 1: is_covid = 0 and Group 2: is_covid = 1

Hypothesis Testing

Null Hypothesis (H_0): mean(is_covid = 0) = mean(is_covid = 1)

Alternative Hypothesis (H_1): mean(is_covid = 0) != mean(is_covid = 1)

Significance Level (α): 5%



Created Binary Indicator

is_covid = 1, if incident date falls between March 2020 and June 2021, 0 otherwise

Statistical Tests

T-Test: Normal distribution, Equal variance between groups.

Mann-Whitney U Test: Non-normal distribution or unequal variance.

Results

Fail to Reject H_0: Feature is unaffected by COVID

Reject H_0: Feature is affected by COVID, added is_covid as an exogenous variable for forecasting

4. Forecasting Independent Features

Primary Model (SARIMAX)

Modeled 124 selected features using SARIMAX, leveraging yearly seasonality with hyperparameter tuning.

Evaluating Forecasts

Calculated MAPE to identify poorly performing forecasts.

Alternative Models

Used Poisson and ETS (Holt-Winters) to get alternative forecast for high-MAPE features.

Model Selection

Chose the model with minimum MAPE for each feature to create final data with forecasted values (2006 - 2034) for each station

Performance Metrics

Performance Metrics Across Stations for Forecasting Independent Features

Metric	Value
RMSE	5.835
MAE	4.740
MAPE	18.945
AIC	743.375
BIC	766.315

Forecasting Accuracy: Metrics demonstrate reasonable forecasting performance, with low errors and deviations across stations.

Model Efficiency: Low AIC (743.375) and BIC (766.315) values highlight a balance between model accuracy and simplicity, avoiding overfitting.

5. Predicting Monthly Counts

Key Input

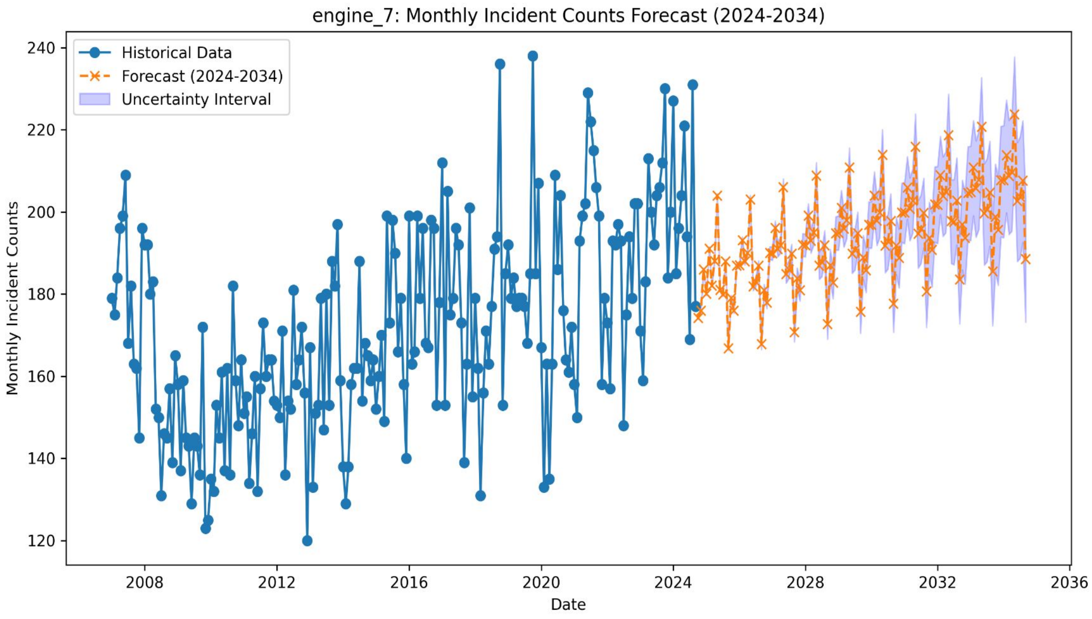
Historical and Forecasted values of important features for each station

Model Used

Prophet Model: Predict monthly incident counts station-wise for the next 10 years

Process

Added forecasted feature values as regressors to the model and predicted monthly counts for each fire station from Oct 24 to Sept 34

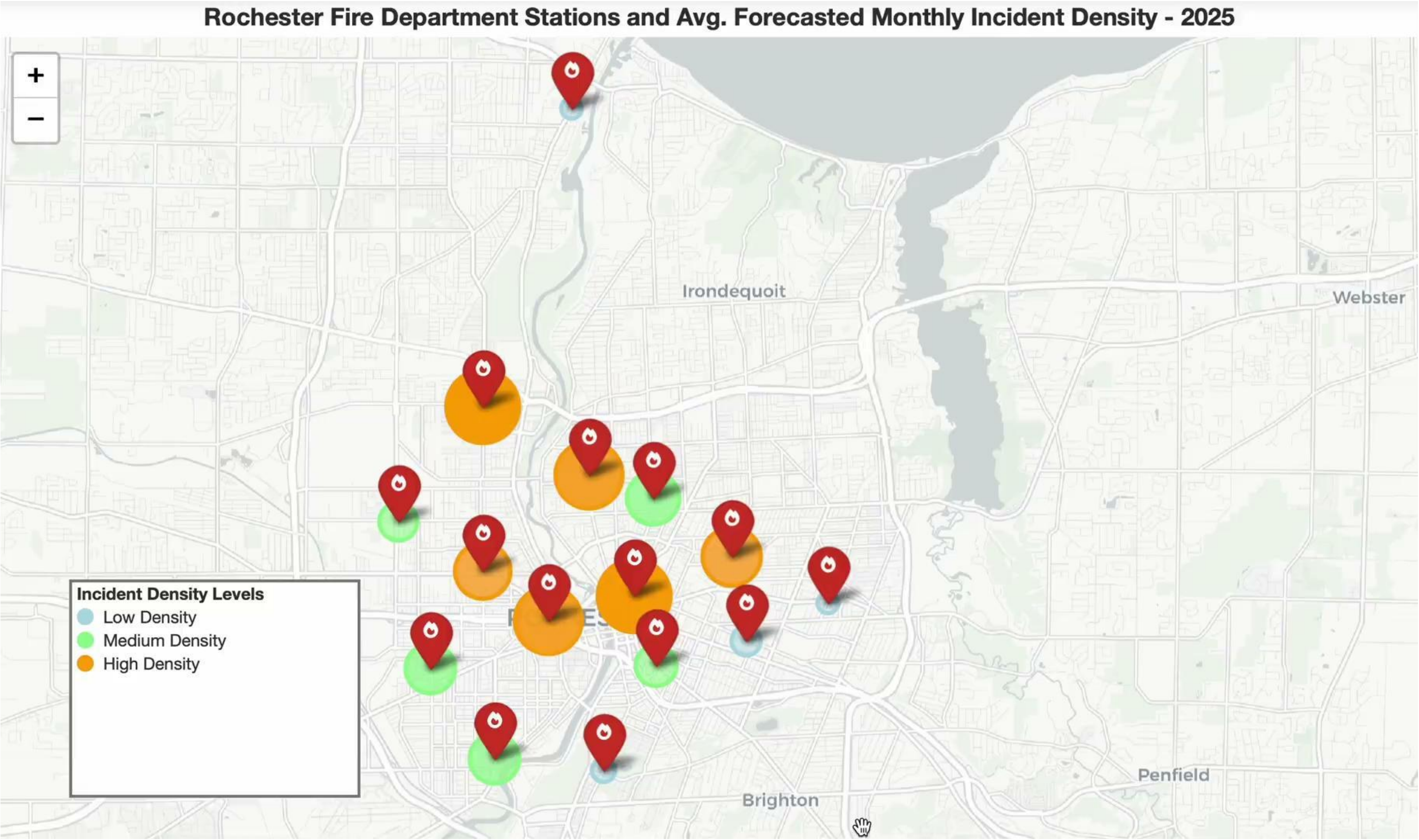


Model Performance

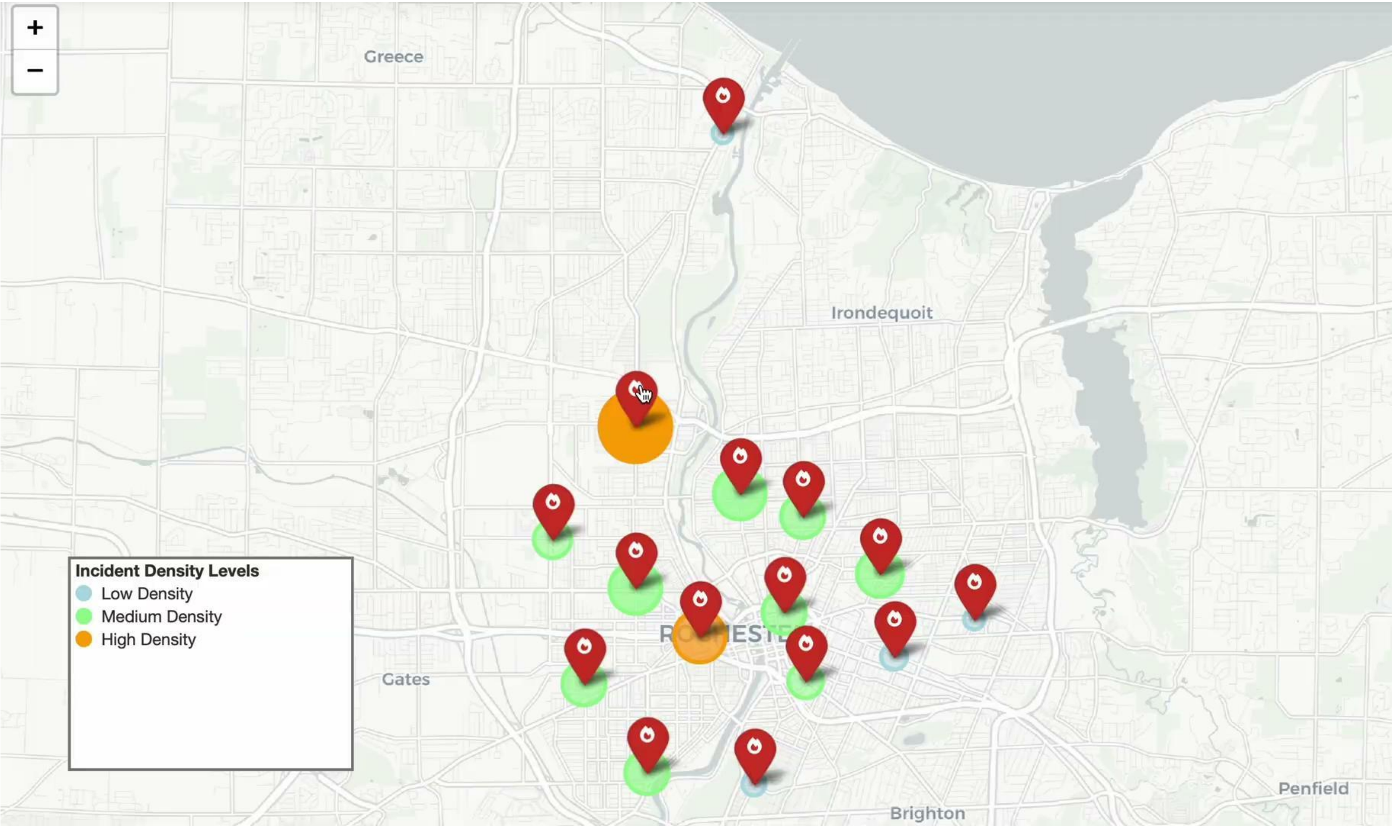
- The Prophet model achieves an average **MAE of 0.34** and **MAPE of 19.2%**, indicating good predictive accuracy across stations.
- An average **RMSE of 0.43** suggests that the model captures trends effectively, with relatively low prediction errors.

station_name	MAE	RMSE	MAPE
engine_1	0.292	0.460	0.142
engine_2	0.782	0.821	0.202
engine_3	0.140	0.190	0.078
engine_5	0.474	0.489	0.261
engine_7	0.184	0.209	0.092
engine_8	0.014	0.017	0.036
engine_9	0.295	0.402	0.108
engine_10_truck_2	0.643	0.671	0.223
engine_12	0.345	0.752	0.535
engine_13_truck_10	0.391	0.471	0.159
engine_16_truck_6	0.373	0.471	0.166
engine_17_rescue_11	0.488	0.570	0.175
engine_19	0.037	0.050	0.053
truck_3	0.133	0.150	0.145
truck_4	0.408	0.528	0.472
truck_5	0.436	0.621	0.218

Map - Avg. Forecasted Monthly Incident Density per Station - 2025



Map - Avg. Forecasted Monthly Incident Density per Station - 2033



Challenges

Integrating Domain Knowledge

Ongoing collaboration and iterative adjustments

External Datasets

Lack of external data to augment the analysis

Poor Predictive Performance for Time to Reach

16 models tested, R^2 score < 30% suggesting weak relationships

Future work

- Explore innovative strategies for better resource use
- Interactive Map Enhancements
- Comprehensive Reporting
- Stakeholder Presentation

Acknowledgements

- **Rochester Fire Department:**
  - Captain Daniel Curran, Head of RFD planning and research
  - Willis Sontheimer, Data Analyst
  - Na Liu, Data Analyst

- **Academic Supervisors**
  - Cantay Caliskan, Professor
  - Ajay Anand, Professor

THANK YOU


Hudson AV
ENG16INE

FIRE DEP

ROCHESTER FIRE DEPT.

ISO CLASS 1
ROCHESTER
ENGINE
16

ROCHESTER


ROCHESTER

ROCHEST